



## L'annuaire du Collège de France

Cours et travaux

112 | 2013

Annuaire du Collège de France 2011-2012

---

# Stochastic models for the inference of live evolution

(CIRB)

Amaury Lambert

---



### Édition électronique

URL : <https://journals.openedition.org/annuaire-cdf/1097>

DOI : 10.4000/annuaire-cdf.1097

ISBN : 978-2-7226-0325-7

ISSN : 2109-9227

### Éditeur

Collège de France

### Édition imprimée

Date de publication : 1 avril 2013

Pagination : 907-909

ISBN : 978-2-7226-0198-7

ISSN : 0069-5580

### Référence électronique

Amaury Lambert, « Stochastic models for the inference of live evolution », *L'annuaire du Collège de France* [En ligne], 112 | 2013, mis en ligne le 22 novembre 2013, consulté le 20 mai 2021. URL : <http://journals.openedition.org/annuaire-cdf/1097> ; DOI : <https://doi.org/10.4000/annuaire-cdf.1097>

---

Collège de France

[P14] Desroches M., Rodrigues S., Krupa M., « Inflection, canards and excitability threshold in neuronal models », *J. Math. Biol.*, 2012, à paraître.

[P15] Krupa M., Vidal A., Desroches M. et Clement F., « Multiscale analysis of mixed-mode oscillations in a phantom bursting model », *SIAM J. Appl. Dyn. Sys.*, 2012, à paraître.

[P16] Touboul J., « Limits and dynamics of stochastic neuronal networks with random delays », *J. Stat. Phys.* sous presse.

---

## Stochastic models for the inference of live evolution (CIRB)

Responsable : Amaury LAMBERT

L'équipe SMILE (*Stochastic Models for the Inference of Life Evolution*; cf. <http://www.proba.jussieu.fr/~smile/>) regroupe des mathématiciens, des biologistes et des bio-informaticiens. Les recherches de notre équipe portent principalement sur la modélisation probabiliste de l'évolution. Durant l'année 2011-2012, nous avons notamment fait des contributions importantes, dans le sens croissant des échelles de temps et d'espace, à la génétique des populations hors d'équilibre, à la reconstruction phylogénétique, et à la modélisation de la diversification des espèces.

## RECHERCHE

### Génétique des populations hors d'équilibre

Les travaux d'Amaury Lambert en collaboration avec Nicolas Champagnat (INRIA) et Lea Popovic (univ. Montréal) sont basés sur la notion de processus ponctuel de coalescence, qui décrit la généalogie à temps fixe d'une population en croissance libre, pour certains processus de branchement non markoviens appelés « *splitting trees* ». Cet objet, mis en évidence par A. Lambert en 2010, est dû originellement à Aldous et Popovic dans le cas particulier des processus de naissance et de mort et de l'arbre brownien. Dans une collaboration avec N. Champagnat, nous utilisons cet objet pour caractériser l'extrémité gauche (familles clonales de taille fixée, par opposition aux plus grandes familles clonales, dont la taille croît avec la taille totale de la population) du spectre de fréquence par allèles pour des *splitting trees* avec mutations poissonniennes, sous un modèle à une infinité d'allèles. Dans la collaboration avec L. Popovic, nous généralisons cet objet à tous les processus de branchement markoviens (en particulier ceux à espace d'états continu), tâche qui nécessite de prendre en compte les points de branchement multiples.

Un article de survol autour de la partition allélique des processus de branchement, à temps fixe ou en régime asymptotique (temps et taille de population grands) vient d'être accepté pour publication.

L'utilisation de ces modèles comme modèles microscopiques neutralistes de la diversification (voir section 3), où la spéciation est le produit de la différenciation génétique, ont donné lieu à des résultats remarquables obtenus par un stagiaire de master de l'ENS (Marc Manceau), et qui seront publiés l'an prochain.

## Reconstruction phylogénétique

Dans une collaboration publiée dans *Molecular Ecology* de deux membres de l'équipe (Amaury Lambert et Guillaume Achaz) avec N. Puillandre (MNHN) et S. Brouillet (UPMC), nous avons proposé une procédure de classification statistique récursive de séquences génétiques homologues en groupes homogènes (espèces) par estimation paramétrique d'un seuil séparant les distances génétiques intra-spécifique et interspécifique. Ce seuil, appelé « *barcode gap* », est ici autorisé à varier en fonction des sous-clades. Notre méthode, basée sur la théorie du coalescent de Kingman, peut être utilisée pour construire le répertoire « *barcode of life* » en classifiant en moins d'une minute des alignements allant jusqu'à plusieurs milliers de séquences homologues.

Dans un travail publié dans *Systematic Biology*, nous (Avner Bar-Hen et co-auteurs) avons étudié l'influence de l'échantillonnage de certains taxons particuliers (homoplasie moléculaire, sélection récente, etc.) sur le résultat de la reconstruction d'un arbre phylogénétique. Nous avons proposé un indice qui quantifie l'influence de chaque taxon sur la reconstruction phylogénétique, puis montré que les taxons influents sont rares mais peuvent bouleverser la reconstruction, même dans des parties éloignées de l'arbre.

## Processus micro et macro-évolutifs de diversification

Deux doctorants de A. Lambert, Robin Aguilée (biologiste) et Mathieu Richard (mathématicien) ont étudié des modèles microscopiques de la diversification, c'est-à-dire basés sur la modélisation des comportements écologiques à l'échelle individuelle. Dans un travail publié dans *Journal of Evolutionary Biology*, nous (A. Lambert et co-auteurs) avons étudié l'effet de la dynamique de la structuration spatiale sur la spéciation, et en particulier comment la spéciation par adaptation locale commandée par de nombreux *loci* en héritabilité mendélienne, peut néanmoins être rendue possible par renforcement après succession de divers stades allopatriques. Nous (M. Richard) avons étudié le *pattern* de biodiversité prédit par un modèle microscopique insulaire où toute la diversité provient du continent par immigration, et montré que ce *pattern* est robuste aux hypothèses classiques en modélisation probabiliste, comme la propriété de Markov de la dynamique de population.

Notre équipe a également beaucoup contribué dans le domaine de la modélisation neutraliste des processus macro-évolutifs. Une de nos publications (Hélène Morlon Todd Parsons et co-auteur), parue dans les *PNAS*, était consacrée au calcul de la vraisemblance d'une phylogénie dans un modèle macro-évolutif de spéciation et extinction dont les taux varient dans le temps, et à l'application de cette méthode à la phylogénie des cétacés. Nous avons en particulier montré que les cétacés ont très probablement subi une radiation explosive suivie par une chute brutale des taux de diversification.

## CONFÉRENCES

« Branching Structures, a Bath-Paris meeting », 15-16 septembre 2011, Paris, intervention de A. Lambert.

« SMPGD 2012, Statistical Methods for Post-Genomic Data », 26-27 janvier 2012, Lyon, Special session, intervention d'A. Lambert.

Conférence de Jacques Monod, « Développements théoriques et empiriques en génomique évolutive », 1-4 avril 2012, Roscoff, organisation par G. Achaz.

2<sup>e</sup> journée *Darwin*, 11 mai 2012, Paris, intervention de A. Lambert

« Mathematical and Computational Evolutionary Biology », 18-22 juin 2012, Montpellier, intervention de A. Lambert

La conférence internationale du projet ANR MANEGE (Modèles aléatoires en écologie, génétique, évolution), dont l'équipe SMILE fait partie intégrante, a eu lieu du 11 au 15 juin 2012 au CIRM (Centre international de rencontres mathématiques, Luminy) sous le titre : « Probability, Population Genetics & Evolution », intervention de A. Lambert.

L'équipe SMILE a organisé au Collège de France une conférence ayant réuni une centaine de participants les 22 et 23 octobre 2012. Cette conférence, intitulée « Phylogenetic approaches to diversification » a été unanimement appréciée des participants.

#### SÉLECTION DE PUBLICATIONS 2011-2012

Aguilée R., Lambert A. et Claessen D., « Ecological speciation in dynamic landscapes », *J. Evol. Biol.*, 24, 2011, 2663-2677.

Bel. L., Bar-Hen A., Petit R. et Cheddadi R., « Spatio-temporal functional regression on paleo-ecological data », *J. Appl. Stat.*, 38(4), 2011, 695-704.

Champagnat N. et Lambert A., « Splitting trees with neutral Poissonian mutations I: Small families », *Stoch. Proc. Appl.*, 122(3), 2012, 1003-1033.

Hanczar B. et Bar-Hen A., « A new measure of classifier performance for gene expression data », *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2012, sous presse.

Lambert A. et Popovic L., « The coalescent point process of branching trees », *Ann. Appl. Probab.*, 2012, sous presse.

Lambert A. « Species abundance distributions in neutral models with immigration or mutation and general lifetimes », *J. Math. Biol.*, 63, 2011, 57-72.

Lambert A. et Trapman P., « Splitting trees stopped when the first clock rings and Vervaat's transformation », *Adv. Appl. Prob.*, 2012, sous presse.

Mariadassou M., Bar-Hen A. et Kishino H., « Taxon influence: Assessing taxon-induced incongruities in phylogenetic inference », *Systematic Biology*, 61, 2012, 337-345.

Morlon H., Parsons T.L. et Plotkin J., « Reconciling molecular phylogenies with the fossil record », *Proceedings of the National Academy of Sciences*, 108, 2011, 16327-16332.

Puillandre N., Lambert A., Brouillet S. et Achaz G., « ABGD, automatic barcode gap discovery for primary species delimitation », *Mol. Ecol.*, 21(8), 2011, 1864-1877.

Richard M., « Limit theorems for supercritical age-dependent branching processes with neutral immigration », *Adv. in Appl. Probab.* 2011, 43, 276-300.

Richard M., « Lévy process conditioned by its height process », *Ann. Inst. Henri Poincaré Probab. Stat.*, 2012, sous presse.

Tully T. et Lambert A., « The evolution of post-reproductive lifespan as an insurance against developmental indeterminacy », *Evolution*, 65, 2011, 3013-3020.